Neoadjuvant rectal score as individual-level surrogate for disease-free survival in rectal cancer in the CAO/ARO/AIO-04 randomized phase III trial

Fokas, E ; Fietkau, R ; Hartmann, A ; Hohenberger, W ; Grützmann, R ; Ghadimi, M ; Liersch, T ; Ströbel, P ; Grabenbauer, G G ; Graeven, U ; Hofheinz, R-D ; Köhne, C-H ; Wittekind, C ; Sauer, R ; Kaufmann, Marco ; Hothorn, Torsten ; Rödel, C

Abstract: Background Surrogate end points in rectal cancer after preoperative chemoradiation are lacking as their statistical validation poses major challenges, including confirmation based on large phase III trials. We examined the prognostic role and individual-level surrogacy of neoadjuvant rectal (NAR) score that incorporates weighted cT, ypT and ypN categories for disease-free survival (DFS) in 1191 patients with rectal carcinoma treated within the CAO/ARO/AIO-04 phase III trial. Patients and methods Cox regression models adjusted for treatment arm, resection status, and NAR score were used in multivariable analysis. The four Prentice criteria (PC1-4) were used to assess individual-level surrogacy of NAR for DFS. Results After a median follow-up of 50 months, the addition of oxaliplatin to fluorouracil-based chemoradiotherapy (CRT) significantly improved 3-year DFS [75.9% (95% confidence interval [CI] 72.30% to 79.50%) versus 71.3% (95% CI 67.60% to 74.90%); P = 0.034; PC 1) and resulted in a shift toward lower NAR groups (P = 0.034, PC 2) compared with fluorouracil-only CRT. The 3-year DFS was 91.7% (95% CI 88.2% to 95.2%), 81.8% (95% CI 78.4% to 85.1%), and 58.1% (95% CI 52.4% to 63.9%) for low, intermediate, and high NAR score, respectively (P < 0.001; PC 3). NAR score remained an independent prognostic factor for DFS [low versus high NAR: hazard ratio (HR) 4.670; 95% CI 3.106-7.020; P < 0.001; low versus intermediate NAR: HR 1.971; 95% CI 1.303-2.98; P = 0.001] in multivariable analysis. Notwithstanding the inherent methodological difficulty in interpretation of PC 4 to establish surrogacy, the treatment effect on DFS was captured by NAR, supporting satisfaction of individual-level PC 4. Conclusion Our study validates the prognostic role and individual-level surrogacy of NAR score for DFS within a large randomized phase III trial. NAR score could help oncologists to speed up response-adapted therapeutic decision, and further large phase III trial data sets should aim to confirm trial-level surrogacy.

DOI: https://doi.org/10.1093/annonc/mdy143

disease-free survival in rectal cancer in the CAO/ARO/AIO-04 randomized phase III trial. Annals of Oncology, 29(7):1521-1527.
DOI: https://doi.org/10.1093/annonc/mdy143

Neoadjuvant rectal score as individual-level surrogate for disease-free survival in rectal cancer in the CAO/ARO/AIO-04 randomized phase 3 trial

E. Fokas[1,13], R. Fietkau[2], A. Hartmann[3], W. Hohenberger[4], R. Grützmann[4], M. Ghadimi[5], T. Liersch[5], P. Ströbel[6], G. G. Grabenbauer[7], U. Graeven[8], R-D. Hofheinz[9], C-H. Köhne[10], C. Wittekind[11], R. Sauer[2], M. Kaufmann[12], T. Hothorn[12,*], and C. Rödel[1,13*], on behalf of the German Rectal Cancer Study Group

[1]Department of Radiotherapy and Oncology, University of Frankfurt, Germany
[2]Department of Radiation Therapy, University of Erlangen-Nürnberg, Germany
[3]Institute of Pathology, University of Erlangen-Nürnberg, Germany
[4]Department of General and Visceral, University of Erlangen-Nürnberg, Germany
[5]Department of General, Visceral and Pediatric Surgery, University Medical Center Göttingen, Göttingen, Germany
[6]Institute of Pathology, University Medical Center Göttingen, Göttingen, Germany
[7]Department of Radiation Oncology and Radiotherapy, DiaCura & Klinikum Coburg, Germany
[8]Department of Hematology/Oncology, Kliniken Maria Hilf GmbH Mönchengladbach, Germany
[9]Department of Medical Oncology, University Hospital Mannheim, Mannheim, Germany
[10]Department of Medical Oncology, University of Oldenburg, Germany
[11]Institute of Pathology, University of Leipzig, Germany
[12]Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Switzerland
[13]German Cancer Research Center (DKFZ), Heidelberg, Germany; German Cancer Consortium (DKTK), Partner Site: Frankfurt
*joint senior authors

Address for Correspondence
Prof Emmanouil Fokas
Department of Radiotherapy and Oncology,
University of Frankfurt
Theodor-Stern-Kai 7
60590 Frankfurt, Germany
Phone: +49-(0)69-63015130
Fax: +49-(0)69-63015091
E-Mail: emmanouil.fokas@kgu.de

ABSTRACT

BACKGROUND: Surrogate endpoints in rectal cancer after preoperative chemoradiation are lacking as their statistical validation poses major challenges, including confirmation based on large phase 3 trials. We examined the prognostic role and individual-level surrogacy of neoadjuvant rectal (NAR) score that incorporates weighted cT, ypT and ypN categories for disease-free survival (DFS) in 1191 patients with rectal carcinoma treated within the CAO/ARO/AIO-04 phase 3 trial.

PATIENTS AND METHODS: Cox regression models adjusted for treatment arm, resection status, and NAR score were used in multivariable analysis. The four Prentice criteria (PC1-4) were used to assess individual-level surrogacy of NAR for DFS.

RESULTS: After a median follow-up of 50 months, the addition of oxaliplatin to fluorouracil-based chemoradiotherapy (CRT) significantly improved 3-year DFS (75.9% [95% CI 72.30-79.50] vs 71.3% [95% CI 67.60-74.90]; P=0.034; PC 1) and resulted in a shift towards lower NAR groups (P=0.034, PC 2) compared to fluorouracil-only CRT. The 3-year DFS was 91.7% (95% CI, 88.2 95.2), 81.8% (95% CI, 78.4-85.1) and 58.1 (95% CI 52.4-63.9) for low, intermediate and high NAR score, respectively (P<0.001; PC 3). NAR score remained an independent prognostic factor for DFS (low vs high NAR: HR 4.670; 95% CI 3.106-7.020; P<0.001; low vs intermediate NAR: HR 1.971; 95% CI 1.303-2.98; P=0.001) in multivariable analysis. Notwithstanding the inherent methodological difficulty in interpretation of PC 4 to establish surrogacy, the treatment effect on DFS was captured by NAR, supporting satisfaction of individual-level PC4.

CONCLUSION: Our study validates the prognostic role and individual-level surrogacy of NAR score for DFS within a large randomized phase 3 trial. NAR score could help oncologists to speed up response-adapted therapeutic decision, and further large phase 3 trial datasets should aim to confirm trial-level surrogacy.

Keywords: rectal cancer, NAR score, surrogate, DFS, trial, prognosis

KEY MESSAGE

In this secondary analysis in 1191 patients treated within the CAO/ARO/AIO-04 phase 3 trial, the neoadjuvant rectal (NAR) cancer score was a prognostic factor and individual-level surrogate endpoint for disease-free survival in rectal cancer. NAR score can predict treatment effects on the clinical outcome and could help oncologists to speed up response-adapted therapeutic decisions.

INTRODUCTION

Following the implementation of preoperative treatment, it has become apparent that rectal cancer response to chemoradiotherapy (CRT) varies considerably, ranging from complete tumor disappearance to lack of response, or even disease progression [1]. Early variables to assess tumor response, such as downsizing, downstaging, and tumor regression grading (TRG), have been proposed to reflect tumor biology, treatment efficacy and patients' prognosis [2, 3], and may be used as surrogates for disease-free (DFS) and overall survival (OS).

The interest in early surrogate endpoints has grown considerably in oncology trials [4, 5]. Surrogate endpoints are early indirect measures of true clinical endpoints and can decrease the number of patients and time needed to complete a trial, enabling early and less costly assessment of the benefit of experimental treatments [6]. The establishment of surrogate markers poses a challenge as it requires rigorous statistical validation using large trial datasets. Also, whether surrogate endpoints reflect true clinical benefit is discussed controversially [4, 5]. In that context, the neoadjuvant rectal (NAR) score has been recently proposed by the National Surgical Adjuvant Breast and Bowel Project (NSABP), the Radiation Therapy Oncology Group (RTOG), and the Gynecologic Oncology Group (GOG) (NRG) Oncology to serve as a potential surrogate for clinical endpoints in trials testing preoperative treatment in rectal cancer The NAR score was developed on the basis of Valentini's nomograms for OS [7] incorporating a weighted combination of the pre-CRT cT-category, and post-CRT ypT- and ypN-categories and represents a pseudo-continuous variable with 24 possible discrete scores, ranging from 0–100 [8]. In the NSABP R-04 trial, the NAR score was classified as low (NAR<8), intermediate (NAR=8-16), and high (NAR>16) according to the tertiles of the observed scores, and lower NAR score was associated with better OS [8].

We investigated the prognostic value of NAR score in the CAO/ARO/AIO-04 randomized phase 3 trial. In that trial, the addition of oxaliplatin to 5-fluorouracil (5FU)-based preoperative CRT resulted in a significant improvement of the primary endpoint, DFS, compared to the standard arm [9, 10]. The aims of the present work were (a) to examine the prognostic role of the NAR score, and (b) to assess whether the NAR score constituted an individual-level surrogate for DFS according to the Prentice criteria (PC) that had been used to confirm that the treatment effect on the surrogate endpoint reflects the treatment effect on the clinical endpoint [11].

PATIENTS AND METHODS

Study design and participants

The CAO/ARO/AIO-04 trial (ClinicalTrials.gov, number NCT00349076) was a multicenter, open-label, two arm randomized phase 3 study. The design, treatment plan and clinical outcome have been described before [9, 10]. The trial received approval by the ethics committee of the University of Erlangen, Germany. A +description of the design, pretreatment and pathologic examination, and follow-up is shown in **Supplementary Methods; Supplementary Figure 1** illustrates the treatment plan. The full trial protocol is provided as **Supplementary Material.**

Neoadjuvant rectal score

The neoadjuvant rectal score (NAR) incorporates cT to account for tumor downstaging, and ypT and ypN that are influenced directly by preoperative treatment [8]. The NAR formula is as follows: $NAR = [5\,pN\text{-}3(cT\text{-}pT)+12]^2 / 9.61$, where cT in $\{1, 2, 3, 4\}$, pT in $\{0, 1, 2, 3, 4\}$ and pN in $\{0, 1, 2\}$. NAR consists of 24 distinct scores that range from 0 to 100. For ypT-category and ypN-category, a relative weight of 3 and 5 was suggested to reflect the impact

of these variables, based on the nomogram of Valentini [7]. The constant 12 is included to maintain all scores inside the brackets as positive. The scaling factor 9.61 was introduced to ensure that the final scores range from 0 to 100.

Statistical analysis

The statistical analyses are described in detail in **Supplementary Methods.**

RESULTS

Patient characteristics and association of NAR with clinicopathologic factors

Between July 2006 and February 2010, 1265 patients were recruited in the trial (CONSORT, **Supplementary Figure 2**). A total of 607 patients were actually treated with fluorouracil plus oxaliplatin-based preoperative CRT (5-FU/OX-CRT), and 625 patients actually received fluorouracil alone during CRT (5-FU-CRT). The NAR score was available in a total of 1191 patients and **Table 1** shows the results after both treatments. 5-FU/OX-CRT led to a statistically significant shift towards lower NAR scores as compared with 5-FU-CRT (P=0.034). The sample tertiles in the CAO/ARO/AIO-04 study population were similar to the NAR cut-offs of 8 and 16, as reported in the NSABP R-07 trial (**Supplementary Table 1**).

Regarding pretreatment clinicopathologic characteristics **(Supplementary Table 2),** the proportion of patients with higher NAR scores was significantly increased in patients with age ≤ median, cN+, and less differentiated tumors. Also, NAR was significantly associated with several pathologic factors after preoperative CRT and surgery, including completeness of surgical resection, ypT-category, ypN-category, pathologic UICC-stage, circumferential resection margin involvement (CRM+) and a longer median interval between completion of preoperative CRT and surgery **(Supplementary Table 3)**. The latter reflects the shifting towards lower ypT-categories and is in accordance to previous studies showing that longer

waiting periods between CRT completion and surgery led to increased tumor regression [12, 13].

The prognostic role of NAR for clinical outcomes

The median follow-up was 50 months (interquartile range=38-61 months). **(Table 2).** In univariate analysis, **(Table 2; Figure 1)** lower NAR score was significantly associated with better 3-year cumulative incidence of DFS, local recurrence, distant metastasis and OS (all P<0.001). We examined the prognostic significance of treatment arms and the clinicopathologic parameters in univariate analysis **(Table 2).** A significantly improved DFS (P=0.034) and local control (P=0.020) were observed following addition of oxaliplatin to 5-FU-based CRT. Patients with complete surgical resection had significantly better DFS (P=0.014) and OS (P<0.001), whereas pathologic stage correlated with all four clinical endpoints (P<0.001 in each case). Circumferential resection margin showed statistical significance for DFS (P<0.001), cumulative incidence of distant metastasis (P<0.001), local recurrence (P<.001) and OS (P= .036). Older patients had significantly worse OS (P=0.001).

We next conducted a multivariable analysis for all four clinical endpoints **(Table 3).** Due to multicollinearity that could lead to statistical bias, NAR and pathologic UICC stage could not be tested within the same model. Also, we excluded CRM as several cases were either missing or unknown. Low vs high NAR score constituted an independent prognostic factor for DFS (P<0.001), the cumulative incidence of local recurrence (P=0.002), the cumulative incidence of distant metastases (P< 0.001) and OS (P<0.001). Similar significant findings were observed for low vs intermediate NAR score with regard to all clinical endpoints with the exception of local recurrence (P=0.068) in multivariable analysis. Complete resection (R0) predicted for better DFS (P <0.001), cumulative incidence of distant metastases

(P=0.005) and OS (P<0.001). The experimental treatment arm was associated with better local control (P=0.021) **(Table 3).**

NAR as a surrogate marker for DFS

We evaluated the surrogacy of NAR score for DFS at an individual-patient level based on the four Prentice criteria [11]: PC 1 (significant treatment effect on DFS, P=0.034; **Table 2**), PC 2 (significant impact of treatment arm on NAR, P=0.034, **Table 1**), and PC 3 (significant association between NAR score and DFS, P<0.001; **Table 2**) were fulfilled. PC 4 necessitates that the significant effect of the treatment arms on the primary endpoint DFS disappears once the surrogate is accounted for. The assessment for PC 4 was based on a Cox model for DFS with treatment arm and NAR included. The previously significant treatment effect on DFS (P = 0.034) as shown in **Table 2** has now vanished from this model (HR 0.880, 95% CI 0.693-1.116, P = 0.292), while the significant impact of NAR score (low vs. high) on DFS was retained (HR = 3.855; 95% CI 1.879-7.909; P < 0.001). Therefore, the treatment effect on DFS was captured by the NAR score, satisfying PC4. Additional sensitivity analyses were conducted to demonstrate the surrogacy of NAR for DFS based on PC 4 (**Supplementary Methods and Results; Supplementary Figure 3**) [14].

DISCUSSION

The NAR score has been proposed by the NRG Oncology as a primary endpoint to assess preoperative treatment efficacy in clinical trials in rectal cancer [8, 15]. The NAR score was prognostic for OS in a retrospective series [16] and the NSABP-R04 phase 3 trial dataset for the entire study cohort but analysis according to treatment arms was not performed [8]. The basic hypothesis is that changes in mean NAR scores between neoadjuvant treatment interventions should translate to changes in DFS or OS. The NAR score retained an

independent prognostic value for the primary endpoint, DFS, in multivariable analysis in the CAO/ARO/AIO-04 randomized phase 3 trial.

Appropriate surrogate endpoints in trials depend on the clinical context, and require careful interpretation [17]. Surrogate endpoints in rectal cancer are lacking as their statistical validation poses major challenges, including confirmation in large trial datasets. In our study, the four PC [11] regarding the individual-level surrogacy of NAR for DFS were met. Several trials have used the PC in the recent years for assessment of potential surrogates [17-21]. It should be noted that the PC 4 is characterized by inherent difficulties in its interpretation that constitutes a methodological limitation in establishing surrogacy, and alternative methods have been proposed by Buyse and colleagues [6, 22, 23]. Also, assessment of NAR effects based on meta-analysis of large randomized trials is a prerequisite for the validation of its surrogacy at both the individual and trial levels [6], extending beyond the PC. Nevertheless, our large phase 3 trial in rectal cancer confirms the individual-level surrogacy of NAR score for DFS and corroborates its use as primary endpoint in (early) clinical trials, such as the NRG-GI002 using radiosensitizers [24], to speed up evaluation of efficacy and access of new treatments.

Other alternative early surrogate endpoints have been proposed, such as downsizing, downstaging, sterilizing lymph nodes, pathologic complete response (pCR), TRG, circumferential resection margin (CRM) involvement, R0 or sphincter sparing resections [25-27]. The strength of the NAR score is the incorporation of both pre- and post-CRT variables to reflect initial tumor extent and tumor response. Weighing these variables based on their relative importance results in 24 possible discrete scores rather than in a dichotomized endpoint, as in the case of downstaging, pCR, CRM, R0. A shift of the pseudo-continuous NAR scores induced by different neoadjuvant interventions likely reflects treatment effects more accurately compared with binary endpoints.

Tumor regression grading has also been proposed to stratify tumor response to CRT and predict prognosis [3, 28, 29] but histopathologic standardization is lacking [30]. Pathologic complete response (pCR) correlates with survival after neoadjuvant treatment in breast cancer and the FDA allowed its use as surrogate endpoint for accelerated approval process [31, 32]. However, in rectal cancer, the role of pCR remains controversial and depends on several factors, such as the dose and the schedule of radiotherapy, combination with chemotherapy, and the time between treatment and surgery [26, 33, 34]. Yothers et al. showed that the NAR score had greater predictive ability than pCR for OS [35]. Finally, parameters such as the quality of total mesorectal excision can affect clinical outcome that could impact NAR and, hence, variability in quality assurance among trials should be considered when assessing the surrogacy of NAR for OS or DFS.

We would like to acknowledge the limitations of our work. First, assessment of the prognostic value and surrogacy of NAR score was done post-hoc. Second, magnetic resonance imaging was not mandatory for baseline staging that could have affected the NAR score, considering the uncertainty of ultrasound when assessing cT- and cN-category. Third, central pathologic review was not conducted. Fourth, confirmation of PC 4 is discussed controversially due to the abovementioned inherent methodological limitations in establishing surrogacy, and alternative methods have been proposed [6, 14, 36] that should be taken into account. Fifth, analyses were done on the individual-level only. Sixth, the NAR score was developed based on Valentini nomogram for OS, whereas in the present work we assessed the surrogacy of NAR for DFS as the latter constituted the primary clinical endpoint. Altogether, our results corroborate the NRG Oncology strategy to use the NAR score as the primary endpoint in early phase rectal cancer trials including induction chemotherapy and molecular therapies. The NAR score constitutes an easily usable endpoint that can predict treatment effects and help oncologists to speed up response-adapted individualized

therapeutic decisions in the era of personalized medicine. These data pave the path for further validation of the NAR score in large phase 3 trial datasets to confirm trial-level surrogacy.

Acknowledgments: Heiko Sülberg, WiSP Research Institute, Langenfeld, Germany, provided on-site monitoring, source data verification, and managed the clinical trial database. We thank the patients, investigators, and institutions involved in this trial.

Funding support: This work was supported by German Cancer Aid (Deutsche Krebshilfe) that provided a research grant for the trial. The funder did not participate or contribute to the trial design, data collection, data analysis, data interpretation, writing the report, or the decision to submit for publication. CR reports grants from German Cancer Aid (Deutsche Krebshilfe), during the conduct of the study. The rest of the authors declare no conflict of interest relevant to the work. No grant numbers applied.

Conflict of interest: We hereby confirm that none of the authors has any conflict of interest relevant to the present work.

REFERENCES

1. Sauer R, Becker H, Hohenberger W et al. Preoperative versus postoperative chemoradiotherapy for rectal cancer. N Engl J Med 2004; 351: 1731-1740.

2. Fokas E, Liersch T, Fietkau R et al. Downstage migration after neoadjuvant chemoradiotherapy for rectal cancer: the reverse of the Will Rogers phenomenon? Cancer 2015; 121: 1724-1727.

3. Fokas E, Liersch T, Fietkau R et al. Tumor regression grading after preoperative chemoradiotherapy for locally advanced rectal carcinoma revisited: updated results of the CAO/ARO/AIO-94 trial. J Clin Oncol 2014; 32: 1554-1562.

4. Buyse M, Sargent DJ, Grothey A et al. Biomarkers and surrogate end points--the challenge of statistical validation. Nat Rev Clin Oncol 2010; 7: 309-317.

5. Saad ED, Paoletti X, Burzykowski T, Buyse M. Precision medicine needs randomized clinical trials. Nat Rev Clin Oncol 2017; 14: 317-323.

6. Buyse M, Molenberghs G, Paoletti X et al. Statistical evaluation of surrogate endpoints with examples from cancer clinical trials. Biom J 2016; 58: 104-132.

7. Valentini V, van Stiphout RG, Lammering G et al. Nomograms for predicting local recurrence, distant metastases, and overall survival for patients with locally advanced rectal cancer on the basis of European randomized clinical trials. J Clin Oncol 2011; 29: 3163-3172.

8. George TJ, Jr., Allegra CJ, Yothers G. Neoadjuvant Rectal (NAR) Score: a New Surrogate Endpoint in Rectal Cancer Clinical Trials. Curr Colorectal Cancer Rep 2015; 11: 275-280.

9. Rodel C, Liersch T, Becker H et al. Preoperative chemoradiotherapy and postoperative chemotherapy with fluorouracil and oxaliplatin versus fluorouracil alone in locally advanced rectal cancer: initial results of the German CAO/ARO/AIO-04 randomised phase 3 trial. Lancet Oncol 2012; 13: 679-687.

10. Rodel C, Graeven U, Fietkau R et al. Oxaliplatin added to fluorouracil-based preoperative chemoradiotherapy and postoperative chemotherapy of locally advanced rectal cancer (the German CAO/ARO/AIO-04 study): final results of the multicentre, open-label, randomised, phase 3 trial. Lancet Oncol 2015; 16: 979-989.

11. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. Stat Med 1989; 8: 431-440.

12. Wolthuis AM, Penninckx F, Haustermans K et al. Impact of interval between neoadjuvant chemoradiotherapy and TME for locally advanced rectal cancer on pathologic response and oncologic outcome. Ann Surg Oncol 2012; 19: 2833-2841.

13. Probst CP, Becerra AZ, Aquina CT et al. Extended Intervals after Neoadjuvant Therapy in Locally Advanced Rectal Cancer: The Key to Improved Tumor Response and Potential Organ Preservation. J Am Coll Surg 2015; 221: 430-440.

14. Heller G. Statistical controversies in clinical research: an initial evaluation of a surrogate end point using a single randomized clinical trial and the Prentice criteria. Ann Oncol 2015; 26: 2012-2016.

15. Czito BG, Deming DA, Jameson GS et al. Safety and tolerability of veliparib combined with capecitabine plus radiotherapy in patients with locally advanced rectal cancer: a phase 1b study. Lancet Gastroenterol Hepatol 2017; 2: 418-426.

16. Raissouni S, Mercer, J., Gresham, G. et al, External validation of the neoadjuvant rectal (NAR) score and Valentini prediction nomogram (VPN): a multicenter study. J Clin Oncol. 2014;32:5 (abstract 3532).

17. Wilson MK, Karakasis K, Oza AM. Outcomes and endpoints in trials of cancer treatment: the past, present, and future. Lancet Oncol 2015; 16: e32-42.

18. Royce TJ, Chen MH, Wu J et al. Surrogate End Points for All-Cause Mortality in Men With Localized Unfavorable-Risk Prostate Cancer Treated With Radiation Therapy vs

Radiation Therapy Plus Androgen Deprivation Therapy: A Secondary Analysis of a Randomized Clinical Trial. JAMA Oncol 2017; 3: 652-658.

19. Scher HI, Heller G, Molina A et al. Circulating tumor cell biomarker panel as an individual-level surrogate for survival in metastatic castration-resistant prostate cancer. J Clin Oncol 2015; 33: 1348-1355.

20. Mauguen A, Pignon JP, Burdett S et al. Surrogate endpoints for overall survival in chemotherapy and radiotherapy trials in operable and locally advanced lung cancer: a re-analysis of meta-analyses of individual patients' data. Lancet Oncol 2013; 14: 619-626.

21. D'Amico AV, Chen MH, de Castro M et al. Surrogate endpoints for prostate cancer-specific mortality after radiotherapy and androgen suppression therapy in men with localised or locally advanced prostate cancer: an analysis of two randomised trials. Lancet Oncol 2012; 13: 189-195.

22. Molenberghs G, Geys H, Buyse M. Evaluation of surrogate endpoints in randomized experiments with mixed discrete and continuous outcomes. Stat Med 2001; 20: 3023-3038.

23. Molenberghs G, Buyse M, Geys H et al. Statistical challenges in the evaluation of surrogate endpoints in randomized trials. Control Clin Trials 2002; 23: 607-625.

24. George TJ YG, Hong TS,McGory Russell M, You YN, Parker W, Jacobs SA, Pogue-Geile KL, Wu CSY, Vijayvergia N, Wolmark N. J Clin Oncol 35(4_suppl):TPS814. Utilizing total neoadjuvant therapy (TNT) in rectal cancer: NRG-GI002, a phase II clinical trial platform.

25. Glynne-Jones R, Mawdsley S, Pearce T, Buyse M. Alternative clinical end points in rectal cancer--are we getting closer? Ann Oncol 2006; 17: 1239-1248.

26. Valentini V, van Stiphout RG, Lammering G et al. Selection of appropriate end-points (pCR vs 2yDFS) for tailoring treatments with prediction models in locally advanced rectal cancer. Radiother Oncol 2015; 114: 302-309.

27. Fokas E, Strobel P, Fietkau R et al. Tumor Regression Grading After Preoperative Chemoradiotherapy as a Prognostic Factor and Individual-Level Surrogate for Disease-Free Survival in Rectal Cancer. J Natl Cancer Inst 2017; 109.

28. Rodel C, Martus P, Papadoupolos T et al. Prognostic significance of tumor regression after preoperative chemoradiotherapy for rectal cancer. J Clin Oncol 2005; 23: 8688-8696.

29. Trakarnsanga A, Gonen M, Shia J et al. Comparison of tumor regression grade systems for locally advanced rectal cancer after multimodality treatment. J Natl Cancer Inst 2014; 106.

30. Chetty R, Gill P, Bateman AC et al. Pathological grading of regression: an International Study Group perspective. J Clin Pathol 2012; 65: 865-866.

31. Cortazar P, Zhang L, Untch M et al. Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. Lancet 2014; 384: 164-172.

32. von Minckwitz G, Untch M, Blohmer JU et al. Definition and impact of pathologic complete response on prognosis after neoadjuvant chemotherapy in various intrinsic breast cancer subtypes. J Clin Oncol 2012; 30: 1796-1804.

33. Bonnetain F, Bosset JF, Gerard JP et al. What is the clinical benefit of preoperative chemoradiotherapy with 5FU/leucovorin for T3-4 rectal cancer in a pooled analysis of EORTC 22921 and FFCD 9203 trials: surrogacy in question? Eur J Cancer 2012; 48: 1781-1790.

34. Maas M, Nelemans PJ, Valentini V et al. Long-term outcome in patients with a pathological complete response after chemoradiation for rectal cancer: a pooled analysis of individual patient data. Lancet Oncol 2010; 11: 835-844.

35. Yothers G GT, Allegra CJ, et al. Predictive validity of NeoAdjuvant Rectal (NAR) Score and pathologic complete response (ypCR) for overall survival (OS) as surrogate

endpoints in rectal cancer clinical trial. Journal of Clinical Oncology 2016; 34, no. 15_suppl: 3533-3. DOI: 10.1200/JCO.2016.34.15_suppl.3533

36. Fleming TR, Powers JH. Biomarkers and surrogate endpoints in clinical trials. Stat Med 2012; 31: 2973-2984.

FIGURE LEGEND

Figure 1. Prognostic significance of NAR score after preoperative chemoradiotherapy and TME surgery in rectal carcinoma. Prognostic significance of NAR for ***(A)*** disease-free survival; ***(B)*** cumulative incidence of local recurrence; ***(C)*** cumulative incidence of distant metastases; and ***(D)*** overall survival. Please note the different numbers at risk shown below each graph, according to the different clinical endpoint definition and available follow-up. Statistical significance was examined using the log-rank test, stratified by treatment arm and the statistical test was two-sided. In univariate analysis, the cumulative incidence of locoregional and distant recurrences was analyzed with death as competing risk, whereas all

Figure 1

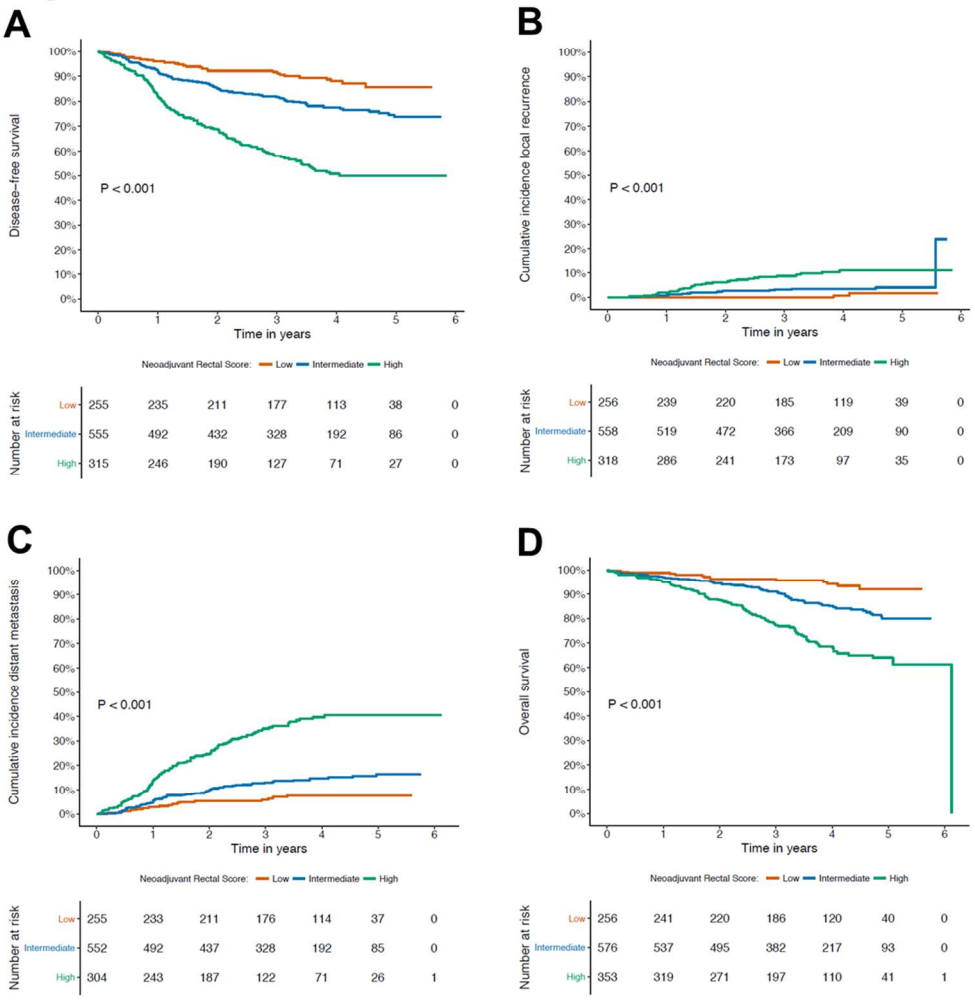

Figure 1

203x216mm (150 x 150 DPI)

Table 1. NAR score in patients treated with preoperative 5-FU based CRT +/- oxaliplatin and surgery

NAR score	Preoperative CRT with 5-FU as received n=625	Preoperative CRT with 5-FU/Ox as received n=607	Total n=1232	P-value
Low	116 (18.6%)	143 (23.6%)	259	**0.034***
Intermediate	305 (48.8%)	273 (45%)	578	
High	188 (30.1%)	166 (27.3%)	354	
Missing	10 (1.6%)	9 (1.5%)	19	
No surgery	6 (1%)	16 (2.6%)	22	

Abbreviations: NAR, neoadjuvant rectal; 5-FU, 5-fluorouracil; CRT, chemoradiotherapy; Ox, oxaliplatin; *p-value was calculated based on continuous NAR scores.

Table 2. Impact of different clinical and pathologic factors on clinical endpoints after preoperative 5-FU+/- Oxaliplatin CRT and surgery

Clinical/Pathologic factors	3-year DFS† (% [95% CI])	P-value	3-year Cumulative incidence of local recurrence after R0/1† (% [95% CI])	P-value	3-year Cumulative incidence of distant metastasis† (% [95% CI])	P-value	3-year OS† (% [95% CI])	P-value
All patients eligible	73.6 (71-76.1)		3.2 (2.1-4.2)		20.6 (18.2-22.9)		88.4 (86.5-90.3)	
Treatment arm (as treated)								
5-FU-CRT	71.3 (67.6-74.9)		3.7 (2.1-5.2)		22.3 (18.9-25.6)		88 (85.4-90.7)	
5-FU/OX-CRT	75.9 (72.3-79.5)	**0.034**	2.6 (1.2-3.9)	**0.023**	18.7 (15.5-22)	0.064	88.8 (86.1-91.4)	0.670
Age								
Age <= median	73 (69.4-76.5)		3.4 (1.8-4.9)		22.6 (19.2-26)		90.7 (88.2-93.1)	
Age > median	74.2 (70.5-77.8)	0.334	2.9 (1.5-4.3)	0.832	18.4 (15.2-21.6)	0.142	86 (83.1-88.9)	**0.001**
Gender								
male	72.9 (69.9-76)		3.7 (2.4-5.1)		21.2 (18.4-24)		88.6 (86.4-90.8)	
female	75.1 (70.4-79.7)	0.310	1.7 (0.3-3.1)	0.150	18.9 (14.7-23.1)	0.280	87.8 (84.2-91.4)	0.746
Completeness of local resection								
R0	78.4 (75.8-80.9)		3.4 (2.3-4.5)		17 (14.7-19.3)		89.8 (87.9-91.6)	
R1	56.1 (33.0-79.2)		6.9 (0-17.6)		23.5 (3.8-43.2)		54.8 (33.8-75.7)	
R2*	0 (0%)		0 (0%)		8.3 (0-25.0)		43.1 (17.1-69)	
Rx	52.7 (25.0-80.5)		0		26.9 (1.8-52)		64.8 (39.6-90.1)	
Unknown/missing	n.a	**0.014**	0	0.265	n.a.	0.534	n.a.	**<0.001**
Pathologic stage								
ypT0N0	93.5 (89.8-97.2)		0 (0-0)		3.2 (0.6-5.9)		96.5 (93.7-99.2)	
ypTisN0	100 (0-100)		0 (0-0)		0 (0-0)		100 (0-100)	
I	89.2 (85.7-92.7)		0.5 (0-1.3)		7.6 (4.6-10.6)		95.5 (93.2-97.9)	
IIA	74 (68.6-79.5)		3.3 (1.1-5.6)		19 (14.2-23.9)		87.9 (83.8-92)	
IIB	64.1 (42.4-85.8)		8.5 (0-21.5)		26.3 (6.7-45.8)		91.3 (79.7-100)	
IIIA	70.4 (60.7-80)		2.8 (0-6.2)		24.4 (15.3-33.5)		90.8 (84.7-96.9)	
IIIB	66.9 (59-74.9)		3.5 (0.3-6.7)		29.5 (21.9-37.2)		86.5 (80.7-92.3)	
IIIC	41.3 (30-52.7)		16.7 (7.8-25.6)		44.3 (32.8-55.8)		68.5 (57.6-79.5)	
IV	0		3.8 (0-9.7)		0		48.2 (33.9-62.5)	
Unknown/Missing	57.1 (30.8-83.5)	**<0.001**	20.3 (0-43.7)	**<0.001**	2.6 (0.2-5)	**<0.001**	61.3 (34.3-88.2)	**<0.001**
Circumferential resection margin								
pCR	93.5 (89.8-97.2)		0 (0-0)		3.2 (0.6-5.9)		96.5 (93.7-99.2)	
<= 1mm	54.9 (39.2-70.6)		19.3 (6.5-32)		31.1 (17-45.2)		69.5 (56.5-82.6)	
> 1mm - 2mm	63.5 (48.1-79)		13.6 (2.9-24.4)		24.5 (10.6-38.3)		73.7 (59.5-87.9)	
> 2mm - Inf	75.5 (72.3-78.7)		2.4 (1.3-3.6)		19.8 (16.8-22.7)		88.3 (85.9-90.7)	
Unknown/missing	80.1 (72.8-87.4)	**<0.001**	4.8 (0.9-8.7)	**<0.001**	13.5 (7.2-19.7)	**<0.001**	87.6 (81.7-93.5)	**0.036**
NAR score								
Low	91.7 (88.2-95.2)		0 (0-0)		5.6 (2.7-8.5)		96.2 (93.8-98.7)	
Intermediate	81.8 (78.4-85.1)		2.7 (1.3-4.1)		13 (10.1-15.9)		91 (88.6-93.5)	
High	58.1 (52.4-63.9)		7.5 (4.5-10.4)		34.5 (29-40)		77.3 (72.6-82)	
Missing (surgery)	90.9 (73.9-100)	**<0.001**	0 (0-0)	**<0.001**	4.5 (0-13.5)	**<0.001**	83.5 (60.4-100)	**<0.001**

The log-rank test was used to calculate statistical significance stratified by treatment arm. The cumulative incidence of locoregional and distant recurrences was assessed considering death as competing risk. The statistical test was two-sided.

Abbreviations: DFS, disease-free survival; OS, overall survival; 5-FU, 5-fluorouracil; CRT, chemoradiotherapy; OX, oxaliplatin; pCR, pathologic complete response; NAR, neoadjuvant rectal; CI, confidence interval; n.a, not applicable;

*For DFS and cumulative incidence of local recurrence, R2 an event.

†With regard to post-surgical pathologic parameters, all clinical endpoints were calculated from date of surgery to prevent length bias.

	DFS†			Cumulative incidence of local recurrences after local R0/R1 resection†			Cumulative incidence of distant metastases†			OS†		
	HR	95% CI	P-value	HR	95% CI	P-value	HR	95% CI	P-value	HR	95% CI	P-value
7	1	—	—	1	—	—	1	—	—	1	—	—
7	0.842	(0.665-1.066)	0.154	0.496	(0.274-0.900)	**0.021**	0.859	(0.653-1.131)	0.280	0.904	(0.678-1.206)	0.493
92	1	—	—	1	—	**—**	1	—	—	1	—	**—**
2	2.650	(1.599-4.391)	**< 0.001**	3.472	(1.359-8.867)	**0.009**	1.835	(0.964-3.495)	0.065	4.957	(3.293-7.460)	**<0.001**
5	1	—	**—**	1	—	—	1	—	—	1	—	**—**
4	1.971	(1.303-2.980)	**0.001**	3.851	(0.903-16.419)	0.068	2.127	(1.256-3.602)	**0.005**	2.388	(1.349-4.225)	**0.003**
5	4.670	(3.106-7.020)	**< 0.001**	10.180	(2.418-42.851)	**0.002**	6.3	(3.775-10.515)	**<0.001**	4.722	(2.694-8.726)	**<0.001**

re conducted using the Cox model for DFS and OS, and the Fine-Gray model for cumulative incidence of locoregional and distant recurrences. The statistical irst fitted Cox (or Fine-Gray) models with treatment arm and completeness of local resection. We then added NAR to this model, but kept the regression arm and completeness of local resection fix (thus, the effect of NAR now describes the additional information contained in NAR). Empty spaces appear as "—" ints.

ase-free survival; OS, overall survival; HR, hazard ratio; CI, confidence interval; 5-FU, 5-fluorouracil; OX, oxaliplatin; CRT, chemoradiotherapy; NAR,

n were also included in the multivariable analysis for cumulative incidence of distant metastases and overall survival.

cal pathologic parameters, all clinical endpoints were calculated from date of surgery to prevent length bias.